Monitoring water surface changes of wetlands using Fuzzy K-Means Algorithm with Sentinel-1 SAR Data

Fifing[1], Hanum Shirotu Nida[2], Adib Hermawan[3]

Indonesia Defense University, IPSC Complex, Bogor 16810, Indonesia

fifing3@gmail.com[1], hanumshirotunida@gmail.com[2], adib_bima@yahoo.co.id[3]

Abstract. Demak Regency is one of the areas on the north coast of Java. As an area located on the coast of Demak, a series of natural disasters, including abrasion and erosion, rising water surfaces by global climate change, land subsidence, have drown some parts of the coastal areas and caused floods that occur regularly. In an effort to increase the resilience of coastal communities in Demak to various disaster problems, it can be done by rehabilitating the green mangrove paths in coastal areas and along the banks of the river. Mangroves play an important role in reducing the impact of climate change, such as reducing the risk of the impact of waves and extreme weather disasters, protecting the coast from abrasion / erosion, inhibiting sea water intrusion and maintaining water quality from pollution threats. Land degradation that occurs on the coast of Demak Regency ranges between 0.06-1.15 meters / year. This causes the elevation in 2025 to decline quite a lot and causes some areas to be below sea level. Research method using Algoritma K-Mean to determine changes in surface water cover of wetlands in Demak Regency in 2014-1019. The results of this study are expected to provide information related to disaster mitigation and can provide input to decision-making from the local government.

1. Introduction

In 1971, it was agreed in a convention on wetlands called the Convention on Wetlands of International Importance, especially as Waterfowl Habitat (Ramsar Convention) [1]. Initially the Ramsar Convention focused on the problems of waterbirds and migratory birds, but in its development it was agreed that wetland conservation was considered more important, and also developed awareness on the integrity of biodiversity and wise aspects of the use of wetlands [2]. The main purpose of this convention is to stop the encroachment and destruction of wetlands, which are fragile and sensitive ecosystems, and their existence is highly depends on how they used and their conservation efforts [3]. Indonesia as the part of world community showed its contribution to the preservation of wetlands by ratifying the convention through Presidential Decree Number 48 year 1991, in commemoration of the World Wetland Day on February 2, 2019. Wetland International Indonesia (WII) conducts a series of celebratory activities with the aim of raising awareness and increasing society’s knowledge about the importance of the preservation of wetlands. Indonesia celebrated the wetland day in Demak Regency with the theme “Introducing the Importance of Mangrove in the Early Stage to Save Our Coastal Areas" [4].

Demak Regency is one of the areas on the north coast of Java. As an area on the coast, Demak caused a series of natural disasters, including abrasion and erosion, rising sea level by global climate change, land subsidence, has drown parts of the coast and causing floods occured regularly. To increase the resilience of the people living on the coast in Demak to various disaster problems, it can be done by rehabilitating the green mangrove paths in coastal areas and along the river bank. Mangroves play an important role in reducing the impact of climate change, such as reducing the risk of catastrophic waves and extreme weather, protecting beaches from abrasion / erosion, inhibiting sea water intrusion and maintaining water quality from pollution threats. Land degradation that occurs on the coast of Demak Regency ranges between 0.06-1.15 meters/year.

This causes the elevation in 2025 quite a lot reduction and causes some areas was below sea level. The sea level tide in 2025 is predicted can reach 1.63 meters and can cause areas with lower elevations inundated by sea water. The area which very high tidal flood hazard reaches 57% of the research area.

The wetlands referred in this study are swamp ecosystems including peat swamps which affect fresh water and surface water on the sea coast. Wetlands include coastal areas, swamplands, peatlands, acidic, natural and synthetic potentially sulphate lands and mangrove areas. The wetland area has several unique characteristics, namely: a. is a lowland along the coast, b. is an area that has low elevation, c. some places are affected by tides for areas which close to the coast, d. influenced by seasons located far from the coast, and e. most of this area is covered with peat. Monitoring water surface of wetlands can utilize remote sensing technology. Remote sensing satellite data with Sentinel 1 Synthetic Aperture Radar (SAR) can be used for monitoring objects on the surface of the earth.

SAR Sentinel 1 data processing has specific advantages in a bad weather during optical data acquisition. SAR data can provide a cloud-free image that effective for observing land cover changes. Technique that can be used for the purpose of monitoring and mapping data is clustering technique. In this study utilizing Sentinel-1 SAR Data that produces a high-resolution satellite image. Clustering techniques can group data based on similarity into clusters. Each cluster has a set of data similar to other data in one cluster, but not similar to the data in other clusters [10]. There are several algorithms that can be used, one of the algorithm that popularly used for clustering a dataset is K-Means. K-Means algorithm is a non-hierarchical clustering method that has relatively fast computing time. Based of the research conducted by Suomi G. and Sanjay Kumar D about the theme of comparative analysis between K-Means and Fuzzy C-Means (FCM) prove that the results of K-Means algorithm is faster with an elapsed time of 0.433755 seconds compared to the FCM algorithm which has elapsed time of 0.781679 seconds [11].

Based on the background of the problems related to environmental problems in Demak Regency such as rising sea levels caused floods, climate change and loss of biodiversity due to the decrease the number of wetland areas in Demak Regency. So, this study discusses changes in surface water cover of wetlands in Demak Regency in 2014-2019. The outcome of this study are expected to provide information related to disaster mitigation and can provide input for decision-making from the local government.

2. Materials dan Methods

2.1 Description of the study area

This research was conducted in the coastal area of Demak Regency, Central Java, especially in the Wedung District area, the image can be seen in **Figure 1**. This study uses sentinel 1 image data in 2015 - 2019 to determine changes in water surface in the wetlands of the coastal area of Demak Regency. The data used in this study are: sentinel-1 SAR images recorded temporally with Vertical Horizontal and Vertical Vertical polarization.

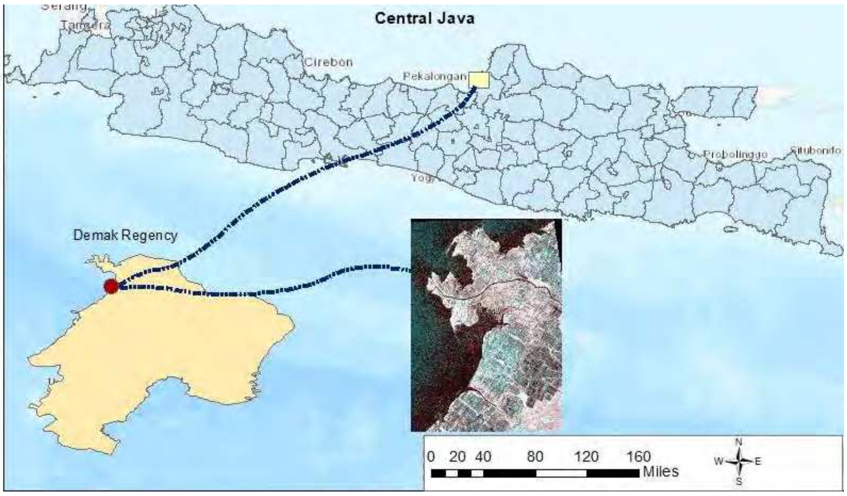


Figure 1. Location of study area in Demak District, Central Java

3. Methods

K-Mean algorithm is an algorithm that requires input parameters as many as k and divides a set of n objects into k clusters so that the level of similarity between members in one cluster is high while the level of similarity with members in other clusters is very low. The similarity of members to clusters is measured by the proximity of the object to the mean value in the cluster or referred to as the centroid cluster [9]. First, K-Mean determines the mean or midpoint of the cluster based on the number of clusters made, then k-mean divides the data set into each of the same clusters or which approach the cluster's mean value [8]. This algorithm aims to minimize an object function. The function of objects in Kmean can be found by equation (1), where $|| x_i^{(j)} - c_j ||$ is the distance of the data point $x_i^{(j)}$ and the center of the C_j cluster is the point n data indicator from the respective cluster center.

$$\sum_{j-1}^{k} \quad \sum_{i-1}^{n} \left\| x_i^{(j)} - c_j \right\|^{2} \qquad (1)$$

K-means algorithm steps: (1) Determine the initial k point on the object to be clustered. (2) Define each object in a group that has the closest objective value. (3) When all objects have been determined, the position of point k is counted again. (4) Steps 2 and 3 are repeated until there are no longer change.

K-Means methods is the most simple and common clustering methods. This is because K-Means have the ability to group large amounts of data with relatively fast and efficient computing time [6]. But this method has a weakness in analyzing the distribution of data and still depends on centroid initialization. K-Means only looks at the distance of data to each centroid on each cluster [7]. Following is an algorithm of the K-Means method:

a. Enter the data to be cloned.

b. Determine the cluster number.

c. Take any data as much as the cluster number randomly as the cluster center (sentroid).

d. Calculate the distance between the data and the cluster center, using the equation

$$D(i,j) = \sqrt{\left(x_{1i} - x_{1j}\right)^2 + \dots + (x_{ki} - x_{kj})^2} \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (2)$$

Where :

D (i,j) = distance of data i to center of cluster j,
x_{ki} = data to i in attribute to j
x_{ji}= center point to j, in attribute k

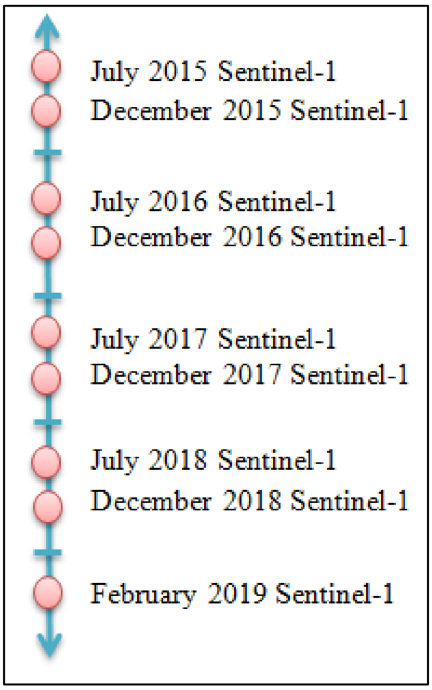
e. Recalculate the cluster center with the new cluster membership
f. If the center of the cluster does not change, the cluster process is complete, if not then repeat step d until the cluster center does not change again.

3.1 SAR Data Pre-Processing

Sentinel 1 product used is Sentinel-1A with IW (Interferometric Wide Swath) format with level 1 GRD data level. Ground Range Detected is data that has been detected and includes multi-looking data. This data has been projected with a field distance using the earth ellipsoid model. The phase information on this product is lost and the specimen is reduced because it reduces its geometric resolution (15). While the IW format is the main acquisition on land. This mode has a coverage of 250 km with a spatial resolution of 5m x 20m single look. IW mode acquired 3 sub-plots using TOPSAR (Terrain Observation with Progressive Scan SAR). TOPSAR mode is intended to replace the conventional ScanSAR mode with the same coverage and resolution but the noise ratio becomes small. This technology has advantages that can produce a homogeneous image quality with noise that can be minimized (Signal-to-Noise Ratio) almost uniformly. Below is shown in Table 3. It is a characteristic of Sentinel 1 SAR Data and in Picture 2 is the time for collecting Sentinel 1 radar images which we discussed in 2015 - 2019.

Tabel 1. Sentinel -1 A data characteristics

Mission	**Sentinel -1A**
Bandwith	5.405 GHz
Wavelength	C (5.6 cm)
Images mode	Interferomentic Wide
Track	153
Orbit	Descending
Product	SLC
Ground resolution	3.1 by 21.7 m
Pixel spacing	2.3 by 13.8 m
Polarization	Dual (VV, VH)
Incidence angle	38.9^0
Revisit time	12 days
Covered area	250 by 170 km



Picture 2. Sattelite image used in the study

The SAR image is actually still difficult to interpret compared to optical images [13] because the intensity of the signal captured by a sensor called the backscatter coefficient (σ °) in decibels (dB) caused by various configurations, various polarizations, various angles resulting in the same surface producing different backscatter coefficient. The interaction between radiation and the internal nature of the tree / surface of the object will produce a certain backscatter signal. Backscatter signal will be influenced by the internal and external components of the object, for example, water content that affects the dielectric constant of material, cell structure, and other external components which are affected by the signal mechanism dissipated, for example the size, geometry and orientation of the leaf, stem, branches, as well as air or stage roots [12].

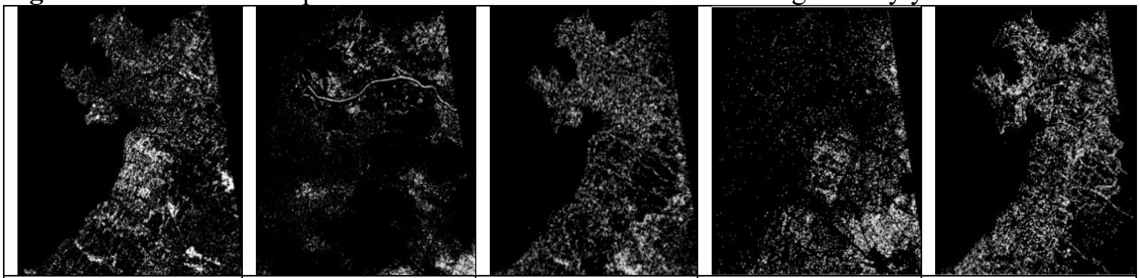
Tabel 2. Polarized indices calculated from Sentinel-1 SAR Data

Index	Abbreviation	Equations	Reference
Polarized Ratio (VH to VV)	VHrVV	$Y^0{}_{VHrVV} = Y^0{}_{VH}/Y^0{}_{VV}$	Brisco et al.[15]
Normalized Difference Polarized Index	NDPI	$Y^0{}_{NDPI} = (Y^0{}_{VV} - Y^0{}_{VH})/\ (Y^0{}_{VV} + Y^0{}_{VH})$	Mitchard et al [16]
Normalized VH Index	NVHI	$Y^0{}_{NVHI} = Y^0{}_{VH}\ /\ (Y^0{}_{VV} + Y^0{}_{VH})$	McNairn & Brisco [17]
Normalized VV Index	NVVI	$Y^0{}_{NVVI} = Y^0{}_{VV}\ /\ (Y^0{}_{VV} + Y^0{}_{VH})$	McNairn & Brisco [17]

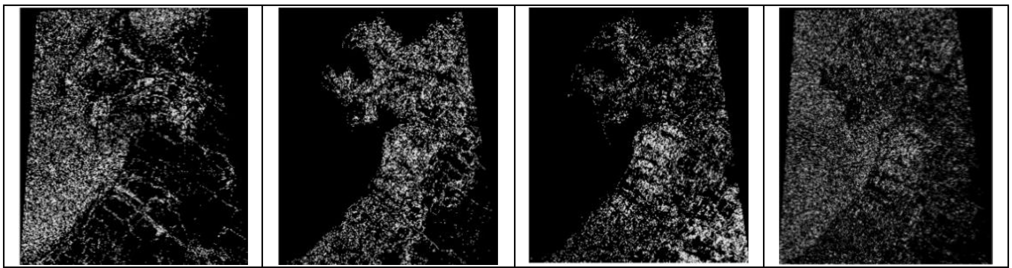
4. Result and Discussion

4.1 Surface Water monitoring With SAR

SAR data is very useful for mapping and monitoring surface water because of its ability to be accessible to the image of the earth's surface without the difference in weather, day and night. The SAR data display is the roughness, texture and hue generated from the backscatter so that visual representation will be difficult to determine the air surface cover in wetlands. The interpretation technique that is carried out to complete is the composite temporal RGB SAR data. Display of **Figure 3.** below shows a pattern of surface distribution which changes every year.



July 2015	Dec 2015	July 2016	Dec 2016	July 2017

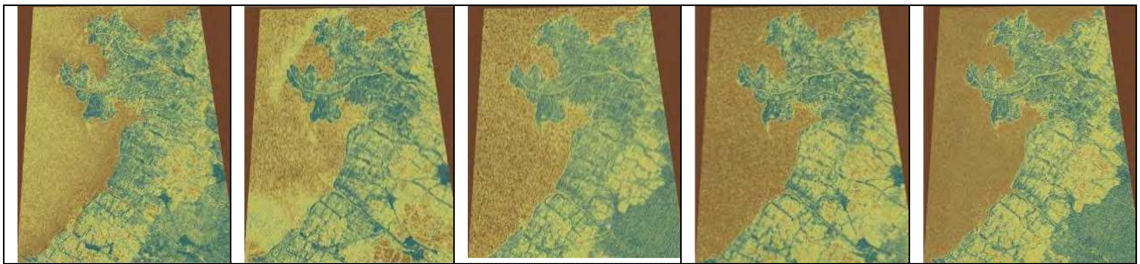


Dec 2017	July 2018	Dec 2018	Mar 2019

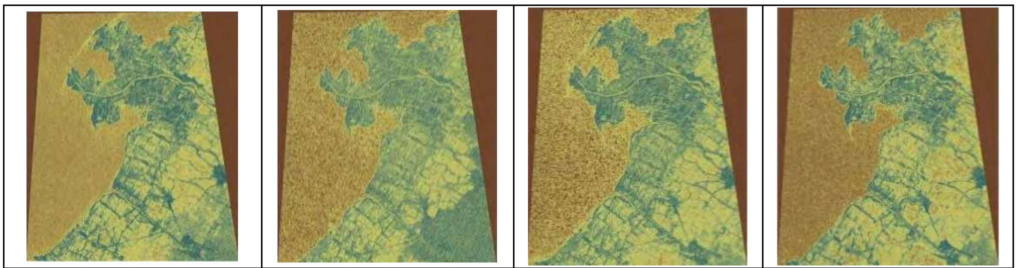
Figure 3. Pattern of distribution of wetland surface water

4.2 Wetland Classification With SAR

The process of classifying land is very difficult to interpret because of various types of wetlands and between several types of wetlands and closure of surface water. This case occurs when using optical recording and electronic wave spectrum. Consequently, using temporal analysis with SAR data sources is the best way. The purpose of this section is to provide information on the use of SAR data in the classification of wetlands. **In figure 4**. You can see composite images from the research area.



July 2015	Dec 2015	**July 2016**	Dec 2016	July 2017



Dec 2017	July 2018	Dec 2018	**Mar 2019**

Figure 4. Sentinel 1 satellite images from 2015 – 2019

4.3 SAR Image Processing for Wetland and Surface Water Mapping

The purpose of unsupervisied classification is to determine the possible number of classes in the wetland area and its surroundings. The results show that there are four wetland classes in this study, namely fields, ponds, mangroves, and open water bodies.

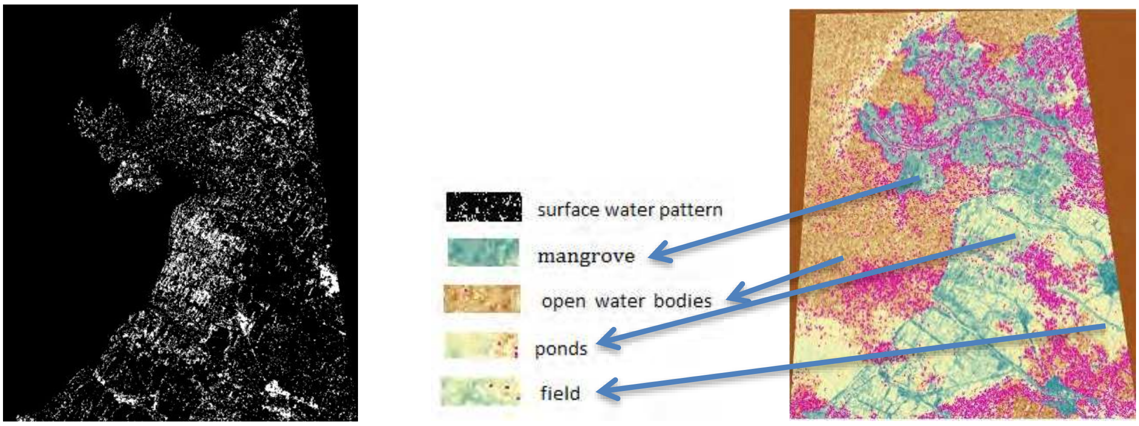


Figure 5. Visual interpretation of multi-temporal water and composit SAR data

Tabel 3. Nilai backscatter pada perubahan air permukaan data multitemporal

Classification	July 2015	July 2016	July 2017	July 2018	March 2019
Field	0.004229	0.008858	0.050254	0.001976	0.178336
Mangrove	0.182700	0.178336	0.001976	0.001976	0.002130
Open water bodies	0.002130	0.000494	0.001976	0.004229	0.182700
Ponds	0.001976	0.178336	0.004229	0.182700	0.001976

SAR visual interpretation Multitemporal water and composit data (can be seen in **Figure 5**) shows that there are four classes of clusters namely mangrove, open water bodies, ponds and fields.

SAR Blue multitemporal RGB composite data indicates the cover of land with the type of mangrove. In the same location SAR data shows the color of dark rice fields which means the backscatter value is low because the reflected signal does not return to the sensor due to the smooth, flat or water surface. This phase is the phase where the rice fields are still irrigated indicating that there is no large vegetation. In July 2018 blue spots began to appear in the same location which indicated the backscatter value received by the sensor. This phase is the phase where wetlands have begun to be overgrown with plants and cause multiple bounces on the radar signal, but black is still visible.

Tabel 4. Classification accuracy assessment

Class	User Accuracy (%)	Commission Error (%)	Producer Accurancy (%)	Omisssion Error (%)
Field	92.3	7.7	85.7	14.3
Mangrove	90.0	10	100	0
Open water bodies	100	0	100	0
Ponds	95.5	4.5	95.4	4.6

In **Table 4.** is classification accuracy assumption, this process is carried out for object classification based on four classes. After the classification of interpretations was made and concluded the highest value of user accuracy, namely open water bodies, the highest Commission error value was mangrove, the highest producer accurancy value was mangrove and open water bodies, while the highest error ommision value was rice fields.

Conclusion

Based on the texture analysis, Sentinel-1A data can be easily monitored for changes in surface water in wetlands. The object analysis process using radar imagery cannot only use one polarization or 2 polarization, but must use two or more polarizations. Based on the comparison of image classification results, it is shown that sentinel-1A imagery can classify wetlands in surface water. Using the k-mean method can classify the wetland class. The K-Mean method can also produce a pattern of distribution of surface water in the wetland cluster. And can provide input to decision-making from the local government.

6. Acknowledgment

This research is supported by The Ministry of Defense of the Republic of Indonesia, Indonesia Defense University, Study Program Disaster Management and Study Program of Sensing Technology. We also would like to thank to our colleagues and lectures for contributing in competing this paper.

References

[1] Batanjski, Vera, et al., “Critical Legal and Environmental View on The Ramsar Convention in Protection from Invasive Plant Species: An Example of The Southern Pannonia Region”, International Environmental Agreements: Politics, Law and Economics , Springer Nature, Vol.16, Issue 6, 2016

[2] Ansari, Haseeb, et al ., 2011, “Ramsar Convention and Convention on Biological Diversity as Mechanism for Promoting Biodiversity Conservation of Malaysian Wetlands”, Journal of Applied Sciences Research , Vol. 7, No. 13, 2011

[3] Finlayson, Max, et al ., “The Ramsar Convention and Ecosystem Based Approaches to the Wise Use and Sustainable Development of Wetlands”, Journal of International Wildlife Law & Policy, Vol. 14, No.3-4, 2011.

[4]]Rilis Pers. 2019, “Memperkenalkan Pentingnya Mangrove Sejak Dini untuk Menyelamatkan Kawasan Pesisir Kita”. Wetlands International (YLBA)

[5] Alfina, Tahta., Santosa, Budi., dan Barakbah, Ali Ridho. Analisa Perbandingan Metode Hierarchical Clustering, K-means dan Gabungan Keduanya dalam Cluster Data (Studi kasus : Problem Kerja Praktek Jurusan Teknik Industri ITS). JURNAL TEKNIK ITS. 2012; Vol. 1, Pp 521-52

[6] Yusuf, Amad & Tjandrasa, Handayani. Prediksi Nilai Dengan Metode Spectral Clustering dan Clusterwise Regression. Jurnal Simantec. 2012; Vol. 4, No. 1.

[7] Zijun, Z. (2012). K-mean algorithm cluster analysis in data mining.

[8] Muzakir, Ari. Analisa dan Pemanfaatan Algoritma K-Means Clustering Pada Data Nilai Siswa Sebagai Penentuan Penerima. Prosiding Seminar Nasional Aplikasi Sains & Teknologi (SNAST). 2014

[9] J. Han, M. Kamber och J. Pei, Data Mining: Concepts and Techniques, Waltham: Elsevier, Inc, 2012.

[10] Ghosh, S. and S. K. Debey, "Comparative Analysis of K-Means and Fuzzy C-Means Algorithms," vol. 4, no 35-39, 2013

[11] Kuenzer C., Bluemel A., Gebhardt S., Quoc V., T., dan Dech S. (2011). Remote Sensing of Mangrove Ecosystems: A Review. Remote Sens, 3(5), 878-928, doi: 10.3390/rs3050878

[12] Kasischke, E., S.; Melack, J., M. dan Dobson, M., C. (1997). The Use of Imaging Radars for Ecological Applications-A Review. Remote Sensing Environment, 59(2), 141-15

[13] ESA. (2013). Sentinel-1 User Handbook, diunduh 3 Agustus 2017 dari https://sentinel.esa.int/documents/247904/685163/Sentinel-1_User_Handbook

[14] Brisco, B.; Kapfer, M.; Hirose, T.; Tedford, B.; Liu, J. Evaluation of C-band polarization diversity and polarimetry for wetland mapping.Can. J. Remote Sens.2011, 37, 82–92.

[15] Mitchard, E.T.A.; Saatchi, S.S.; White, L.J.T.; Abernethy, K.A.; Jeffery, K.J.; Lewis, S.L.; Collins, M.;

[16] Lefsky, M.A.; Leal, M.E.; Woodhouse, I.H.; et al. Mapping tropical forest biomass with radar and spacebornelidar in lopénational park, gabon: Overcoming problems of high biomass and persistent cloud.Biogeosciences 2012, 9, 179–191. [CrossRef]

[17] McNairn, H.; Brisco, B. The application of C-band polarimetric SAR for agriculture: A review.Can. J. Remote Sens.2004, 30, 525–542.